



Общество с ограниченной ответственностью
«Дата Интегрейшн Софтвер»
125284, Россия, Москва, Ленинградский проспект, д. 31А, стр. 1, БЦ «Монарх»
ИНН 7713555858, ОКПО 77352347, ОКАТО 452777568000
Тел. + 7 (495) 645-0201, Факс + 7 (495) 645-0188
info@dis-group.ru

DATA QUALITY: DATA QUALITY MANAGEMENT FOR THE DEVELOPER

Слушатели данного тренинга получают знания и навыки, необходимые для начала самостоятельной работы с продуктом Informatica Data Quality 10. Материал семинара построен на основании уникальных материалов и методик разработанных вендором.

В рамках тренинга рассмотрены необходимые теоретические знания по продукту. В добавление к этому, участники тренинга научатся очищать, стандартизировать и обогащать данные. Также в рамках тренинга будут рассмотрены вопросы тестирования и поиска ошибок в проектах, построенных на базе данного продукта.

После успешного завершения тренинга слушатели будут:

- Понимать процессы Data Quality;
- Понимать архитектуру Data Quality;
- Чувствовать разницу между ролями Analyst и Developer, и знать средства разработки, необходимые для каждой из этих ролей;
- Знать средство разработки Developer;
- Уметь выполнять действия по стандартизации и очистке данных;
- Уметь выполнять развертывание решения;
- Уметь выполнять диагностику решения.

Тренинг, в первую очередь, будет интересен следующим категориям специалистов:

- разработчикам, непосредственно работающим с продуктом Informatica Data Quality;
- аналитикам и руководителям проектов, менеджерам по качеству, участвующим в проектах по управлению качеством данных;
- кроме того, к дополнительной целевой аудитории также относятся квалифицированные специалисты, желающие понять суть работы систем автоматизирующих процессы стандартизации данных и желающие повысить свой профессиональный уровень.

Продолжительность тренинга составляет 4 дня (32 академических часа)



Общество с ограниченной ответственностью
«Дата Интегрейшн Софтвер»
125284, Россия, Москва, Ленинградский проспект, д. 31А, стр. 1, БЦ «Монарх»
ИНН 7713555858, ОКПО 77352347, ОКАТО 452777568000
Тел. + 7 (495) 645-0201, Факс + 7 (495) 645-0188
info@dis-group.ru

Каждый посетитель семинара должен иметь при себе ноутбук следующей конфигурации:

Компьютер PC или Mac: <ul style="list-style-type: none">• не ниже CPU: 1.3 GHz• не менее 2 GB RAM• Размер экрана не менее 1024 x 768 или выше	Операционная система: <ul style="list-style-type: none">• Windows XP и выше• Mac OS X• Linux
Поддерживаемые браузеры (не ниже): Internet Explorer 10.0, Firefox 15, Safari 6.0, Opera 12.0, Chrome 16	

Для наиболее полного усвоения программы участникам тренинга предварительно рекомендуется иметь опыт работы с СУБД и общее понимание концепции реляционной модели данных. Также предполагается, что участники тренинга имеют представление об алгоритмическом программировании и владеют английским языком на уровне чтения технической документации.



Программа тренинга по курсу DATA QUALITY: DATA QUALITY MANAGEMENT FOR THE DEVELOPER

Модуль 1: Вступление

- *Обзор программы*
- *Общее описание и назначение курса*

Модуль 2: Обзор процесса Data Quality Process

- *Цикличность процесса управления качеством*
- *Количественные показатели для оценки качества*
- *Процесс управления качеством*
- *Роли Developer и Analyst Roles, и их инструменты*
- *Архитектура Data Quality*

Модуль 3: Проекты и решения Data Quality

- *Примеры внедрения Data Quality*
- *Примеры успешных проектов при внедрении очистки и стандартизации данных*
- *Data Quality и типичные проекты внедрения/обеспечения качества*
- *Отчетность и обеспечение чистоты данных в проектах*
- *Архитектура проектов с Data Quality*

Модуль 4: Совместная работа в проектах и работа с Reference Table

- *Среда разработки Developer*
- *Общие понятия о проектах в Analyst, Data Objects, Profiles, Rules, Scorecards, Comments и Tags*
- *Reference Tables и процессы Data Quality*
- *Создание Reference Tables*
- *Лабораторная работа: Обзор проекта созданного в Analyst*
- *Лабораторная работа: Создание Reference Tables*

Модуль 5: Работа со средством разработки Developer

- *Tasks в интерфейсе Developer*
- *Работа с физическими и логическими объектами*
- *Подключение к таблицам*
- *Импорт и работа с плоскими файлами*
- *Создание logical data objects*
- *Трансформации в Developer*
- *Маппинги и маплеты*
- *Content sets и их пользователи*
- *Особенности работы с Developer*
- *Лабораторная работа: Создание проекта и выдача прав*
- *Лабораторная работа: Создание подключения к базе данных и импорт из файла*
- *Лабораторная работа: Разработка Logical Data Object*

Модуль 6: Профилирование, маплеты и правила

- *Профилирование полей таблиц*
- *Маплеты и Scorecards*
- *Техника профилирования для отладки и повышения качества разработки*



- *Лабораторная работа: Создание правила для количественной оценки качества данных в поле таблицы.*
- *Лабораторная работа: Использование Informatica Analyst, применение правил и Scorecard и просмотр результатов.*

Модуль 7: Стандартизация, Очистка и Обогащение данных

- *Стандартизация, очистка и обогащение данных*
- *Создание маппингов для очистки, стандартизации и обогащения данных*
- *Разработка маплетов для стандартизации*
- *Настройка трансформаций процесса стандартизации*
- *Лабораторная работа: Разработка маппинга с функцией стандартизации и маплетов использующих трансформацию Standardization*

Модуль 8: Разбор данных

- *Процесс разбора данных*
- *Техника применения процессов разбора данных*
- *Основные трансформации для выполнения разбора данных*
- *Лабораторная работа: Выполнение разбора данных с использованием соответствующих трансформаций*
- *Лабораторная работа: Завершение разработки маппинга, стандартизирующего данные*

Модуль 9: Матчинг данных

- *Определение процесса матчинга*
- *Процесс матчинга в DQ*
- *Стадии матчинга*
- *Группировка и ее влияние на матчинг*
- *Методы группировки*
- *Результаты группировки и их зависимость от выбранной стратегии*
- *Алгоритмы матчинга*
- *Лабораторная работа: Разработка и настройка маппингов с группировкой и матчингом.*

Модуль 10: Ручная обработка ошибок и процесс консолидации

- *Управление ошибочными и дублированными записями*
- *Обработка исключительных ситуаций*
- *Работа с таблицами для хранения ошибочных и дублированных данных.*
- *Лабораторная работа: Разработка маппинга для выявления ошибочных данных.*
- *Лабораторная работа: Разработка маппинга для выявления дублей данных*

Модуль 11: Разработка и развертывание решений

- *Workflows и Workflow Tasks*
- *Human Tasks и Шаги (Steps)*
- *Выявление ошибочных и дублированных строк.*
- *Развертывание и запуск workflows*
- *Проверка Tasks в Informatica Analyst.*
- *Лабораторная работа: Разработка Workflow для заполнения Analyst Inbox с Exception Tasks*
- *Лабораторная работа: Разработка Workflow для наполнения Analyst Inbox с Duplicate Record Tasks*



Общество с ограниченной ответственностью
«Дата Интегрейшн Софтвр»
125284, Россия, Москва, Ленинградский проспект, д. 31А, стр. 1, БЦ «Монарх»
ИНН 7713555858, ОКПО 77352347, ОКАТО 452777568000
Тел. + 7 (495) 645-0201, Факс + 7 (495) 645-0188
info@dis-group.ru

Модуль 12: Развертывание: Исполнение маппингов не средствами Developer

- *Настройки при развертывании*
- *Маппинги как приложения*
- *Постановка маппингов, профилирование и Scorecards на расписание*
- *Лабораторная работа: Постановка маппингов на расписание используя Informatica Scheduler.*

Модуль 13: Импорт и экспорт Объектов проекта

- *Экспорт/импорт - примеры проектов и решений*
- *Базовые и расширенные возможности импорта*
- *Экспорт проекта*
- *Лабораторная работа: Импорт проекта с использованием базовых методов.*
- *Лабораторная работа: Импорт проекта с использованием дополнительных возможностей.*
- *Лабораторная работа: Экспорт проекта.*

Модуль 14: Диагностика

- *Распространенные ошибки*
- *Наиболее часто встречающиеся настройки маппингов и трансформаций*
- *Распространенные ошибки конфигурации Workflow*
- *Советы по работе с средством разработки Developer*
- *Лабораторная работа: (По желанию). Диагностика конфигурации маппинга*